DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

How To Gerrymander: A Formal Analysis

Katerina Sherstyuk

SOCIAL SCIENCE WORKING PAPER 855

July 1993

How To Gerrymander: A Formal Analysis

Katerina Sherstyuk

Abstract

The paper presents an effort to incorporate geographic and other possible exogenous constraints that might be imposed on districting into an optimal partisan gerrymandering scheme. We consider an optimal districting scheme for a party which maximizes the number of districts that it will, in expectation, win, given arbitrary distributions of voters and party supporters over the electoral territory. We show that such a scheme exists if an equal size requirement is the only constraint imposed on districting. If, further, the requirement of territorial connectedness is imposed, the optimal districting scheme still exists when arbitrarily small deviations from the equal size requirement are admissible. Additional constraints imposed on districting make gerrymandering more difficult and sometimes impossible. Although the party is assumed to ignore the risk associated with possible shifts in electoral votes and thus takes the expected share of votes as a perfect predictor of electoral outcomes, the presented approach is valid for a party with any attitude towards risk and for any kind of majority rule used in elections. The results are consistent with earlier findings on unconstrained optimal partisan gerrymandering.

How To Gerrymander: A Formal Analysis

Katerina Sherstyuk*

1 Introduction

While the scientific debate over multiple criteria that should be used to provide for fair districting is neverending among political scientists (Balinski and Young (1982), Grofman, Lijphart, McKay and Scarrow (ed.) (1982), Cain (1984), Grofman (ed.) (1990), Butler and Cain (1992)), in practice contiguity and population equality continue to be the most important requirements for any redistricting process. What possibilities do these requirements leave for deliberate gerrymandering by particular interest groups or political parties? This is the problem that we address in this paper. Taking the viewpoint of a political party in charge of the redistricting process, we ask the question: how should the gerrymandering party proceed, given the restrictions that are imposed on districting? And, further, what can be done to prevent gerrymandering?

We take a formal approach to gerrymandering. While certain gerrymandering techniques such as "concentration gerrymanders" and "dispersal gerrymanders" (Owen and Grofman (1988)) are well understood and commonly recognized in the literature, few studies treat the issue as an optimization problem for the group in control of the redistricting process. The approach we present in this paper is in many respects related to Owen and Grofman (1988), who analyze optimal gerrymandering schemes for a risk-averse party in an uncertain world. They consider two possible cases: one in which a party maximizes its expected seat share, and another where a party maximizes the probability that it will win a legislative majority. They find that the optimal partisan gerrymander in both cases looks much like a bipartisan gerrymander, with one set of districts having majorities for the controlling party (we call them the "winning" districts) and the other concentrating the opposition party supporters (the "losing" districts). Specific characteristics of the winning and losing districts, they find, depend on the type of uncertainty the

*I would like to thank Richard McKelvey, Kim Border, and Morgan Kousser for their help and suggestions. Any errors are my own.

party faces and the party's attitude towards risk. If the gerrymandering party is risk-neutral or faces no uncertainty, an optimal gerrymandering scheme involves spreading the party supporters among a maximal possible number of winning districts which are carried by a bare majority. Increased degree of uncertainty and a higher degree of risk-aversion cause the party to form fewer winning districts with the higher concentration of party supporters and, hence, higher probability of winning.

Defining the "ideal" characteristics of winning and losing districts is the first step in the optimal gerrymandering scheme. Implementing the ideal scheme on a given territory is the second, and often a more difficult one. Geographic constraints and the cotiguity of districts requirement, if imposed, may present considerable obstacles in drawing the map with a set of desirable characteristics. Theoretical possibility of implementing the latter step is the main subject of this paper.

In what follows, we try to incorporate geographic and other exogenous constraints that might be imposed on districting into an optimal gerrymandering scheme, paying special attention to the population equility and contiguity (connectedness) requirements. We consider the possibility of an optimal districting scheme for a party that maximizes the number of expected winning seats for itself in a legislative body, with the distribution of the voters over the territory exogenously given. We show that such a scheme exists if a population equality requirement is the only constraint imposed on districting for any continuous distribution of population. If, further, the requirement of territorial connectedness of every district is present, the optimal districting scheme still exists if arbitrarily small deviations from the equal size requirement are admissible.

We further show that imposition of additional requirements on districting, such as ethnic fairness (in the sense of equality), makes realization of the optimal gerrymandering scheme more difficult and sometimes impossible. Thus, the imposition of multiple criteria for districting might be useful in preventing strategic manipulation of electoral outcomes, even if it does not always guarantee fair representation.

To illustrate the details of the districting procedure, we take the case of a party which is not concerned with possible shifts in electoral votes and thus takes the expected vote share as a perfect predictor of future electoral outcomes. Yet, one could view the procedure we propose more generally, as the one which shows the possibility of implementing any feasible districting scheme on a territory with given characteristics, once the desirable characteristics of the map are determined.

Section 2 presents the general existence theorems, followed in Section 3 by an example of optimal districting for the case of a uniform distribution of voters and a single-peaked distribution of partisans over the territory. Conclusions and possible extensions are presented in Section 4.

2 Optimal districting for arbitrary distributions of voters' characteristics

Consider the problem of a political party (or its agent) which is entitled to divide a given territory $B \subset R^2$ into k voting districts. Assume the party maximizes the number of districts which it will, in expectation, win in the election. Suppose that the number of districts, k, is given and the districts must be equal in population. To state the problem formally, consider a measurable space $(B, \mathcal{B})$, where $\mathcal{B}$ is a σ-algebra of subsets of B. Let μ denote the Lebesgue measure of population size defined on $\mathcal{B}$ such that $\mu(B) = 1$. Then the equality of population requirement imposed on i-th district M_i , $i = 1, .., k$, can be presented in the form:

$$\mu(M_i) = 1/k \ ,$$

and we can introduce the following notions:

Definition 1 *Given a territory $B \subset R^2$ and a number k, a map $M = \{M_1,..,M_k\}$ is a collection of k subsets of B satisfying*[1]

1. $M_i \in \mathcal{B}$ *for all* $i = 1, .., k$;
2. $M_i^\circ \bigcup M_j^\circ = \emptyset$ *for any* $i \neq j,\ i, j = 1, .., k$;
3. $\overline{(\bigcup_{i=1}^{k} M_i)} = B$.

Definition 2 *Let the population equality be the only constraint imposed on districting. Then a feasible map is a map M that satisfies*

$$\mu(M_i) = 1/k \text{ for every } i = 1, .., k \ . \tag{1}$$

Suppose that each point x of the territory B is characterized by the expected share of votes $f(x)$ that the redistricting party will get in the election. Assume that the distribution of voters over the territory as well as the expected vote share $f(x)$, $x \in B$, are continuous, exogenously given and cannot be affected by the way the district lines are drawn. Suppose the simple majority rule is used in the elections within each district. Under these assumptions and if the party is maximizing the number of districts which it will, in expectation, win, how should the party draw a districting map?

Under these assumptions, the problem is similar to the divide-a-cake problem studied by Dubins and Spanier (1961), and specifically to the "problem of the Nile." Dubins and

[1]Hereafter, we use the standard notation $\overline{S}$ and S° to denote the closure and the interior of a set S, respectively.

Spanier present a solution to the problem of partitioning a set into k pieces and then evaluating each of n measures on each piece. In this paper, we apply their approach to the districting problem.

Consider how the redistricting party[2] might reason. The party expects to win the election in a district if its expected share of votes there exceeds one half. It is reasonable to assume that the party will not want the expected share of votes to be much greater than one half, since that would mean "wasting" votes for that party that could help to carry other districts. Formally, we introduce the following definiton.

Definition 3 *A district M_i is called winning if $E(f(x)/M_i) > 1/2$.*

We further assume that the party will want to draw the district lines in a way that will yield the maximum possible number of winning districts. That is,

Definition 4 *For any map M, define the set of the winning districts $I(M)$ as*

$$I(M) = \{i|E(f(x)/M_i) > 1/2\} \ . \tag{2}$$

Then let the seat value of the map $m(M)$ be the number of winning districts, i.e. the number of the elements in the set I:

$$m(M) = |I| \ . \tag{3}$$

Definition 5 *A seat-maximizing map is a feasible map that maximizes the number of winning districts. Let $\{SM\}$ denote the set of all seat-maximizing maps.*

Assumption 1 *(seat-maximization) The party prefers any seat-maximizing map to any map that is not seat-maximizing.*

Together with the requirement of equal district size, this reasoning implies that the party will want to find the largest possible area (with respect to population) A contained in B such that the expected share of the votes over this area exceeds one half: $E(f(x)/A) > 1/2$. This region will be then adjusted and divided into m winning districts in a way that will conform to the imposed equality of population constraint. If the party cares only about the number of districts that it will, in expectation, win, then the

[2] We will refer to the redistricting party hereafter simply as "the party", since the whole analysis is presented solely from its point of view.

problem of finding a seat-maximizing map trivially has a solution since its objective function is integer-valued. However, if the party is not indifferent among all seat-maximizing maps, a special procedure should be desined to construct the most preferred, among seat-maximizing, maps. We impose the following additional assumption on the party's preferences.

Assumption 2 *(lexicographic preferences) Among the set of all seat-maximizing maps, the party prefers a map which solves*

$$\max_{M \in \{SM\}} [\min_{i \in I(M)} E(f(x)/M_i)] \tag{4}$$

Thus, although the party considers every district with the expected share of votes above one half to be winning, it prefers to draw a map in a way that will keep the expectation of votes in the winning districts as high as possible, as long as it does not decrease the number of winning districts. We summarize the party's preferences in the following definition of an optimal map.

Definition 6 *An optimal map $M^* = \{A_1, .., A_k\}$ is a seat-maximizing map that solves (4).*

After the winning districts are drawn, the rest of the territory will be divided into n losing districts. (Note that given the above assumptions, it should not matter for the party how the losing districts are drawn). The number of districts m and n will be determined to meet the equal size requirement, and so that $m + n = k$. Observe that if $m > n$, the party may secure, in expectation, the majority of seats in the legislative body.

We now turn to the first proposition of the paper, which formalizes the above reasoning and proves the existence of an optimal map.

Proposition 1 *There exists an optimal map $M^* = \{A_1, .., A_k\}$.*

Outline of the proof of Proposition 1. Following the reasoning presented above, the optimal districting map can be designed in two steps. In step 1, we need to pick out the region A in B such that, first, the expected vote share over this region slightly exceeds one half, and second, A is the biggest in terms of population among all the regions that satisfy the first property. We further estimate the number m of winning districts by comparing the population in A and $B \setminus A$. In step 2, we partition the region A into m

sets of equal size in a way that preserves the expected share of the votes at the value higher than 1/2 in each set. After the winning districts are defined, we divide the rest of the territory, $B \setminus A$, into $n = (k - m)$ regions of equal size; these regions will constitute the "losing" districts.

To prove that this procedure is implementable, we apply Lyapunov's Theorem on the range of a vector measure, as presented in Hildenbrand (1974):

Theorem (Lyapunov) 1 *Let ν_i $(i = 1, ..., m)$ be atomless measures on a measurable space $(\Omega, \mathcal{B})$. Then the set*

$$\{(\nu_1(E), ..., \nu_m(E)) \in R^m \mid E \in \mathcal{B}\}$$

is a closed and convex subset in R^m.

We as well use the corollary to the Lyapunov theorem which is due to Dubins and Spanier (1961):

Corollary (Dubins and Spanier) 1 *If each ν_i is a nonatomic probability measure, then given k and $a_1, .., a_k \geq 0$ with $\sum a_j = 1$, there exists a partition $A_1, .., A_k$ of a set U such that $\nu_i(A_j) = a_j$ for all $i = 1, .., m$ and $j = 1, .., k$.*

First we present the following lemma[3].

Lemma 1 *An optimal map $M^* = \{A_1, .., A_k\}$, if exists, satisfies*

$$E(f(x)/A_i) = E(f(x)/A_j) \quad \text{for all} \quad i, j \in I(M^*) \tag{5}$$

We now turn to the proof of Proposition 1.

Proof of Proposition 1.

STEP 1. 1.1) Given a mesurable space $(B, \mathcal{B})$, where $B \subseteq R^2$ and $\mathcal{B}$ is a σ-algebra of subsets of B, a Lebesgue measure of population size μ defined on $\mathcal{B}$, and $f(x)$ such that $0 \leq f(x) \leq 1$ for any $x \in B$, find a set $\mathcal{A} = \{A \subseteq B \mid E(f(x)/A) \geq 1/2\}$. Pick A^* such that $\mu(A^*) = \max_{A \in \mathcal{A}}[\mu(A)]$. To see that it is possible to find such a set A^*, observe that if the set $\{x \in B \mid f(x) \geq 1/2\}$ is non-empty, then by Lyapunov Theorem we can find at least one set $A \subseteq B$ such that $E(f(x)/A) \geq 1/2$. Next, defining a measure $\eta \in R^2$

[3] The proof of lemma 1 is given in the appendix.

by $\eta_1(C) = \mu(C)$ and $\eta_2(C) = \int_C f(x)d\mu$, we obtain that by the Lyapunov Theorem the range of $E(f(x)/C) = \eta_2(C)/\eta_1(C), C \subseteq B$, is compact, and so is the range of $E(f(x)/C)$ constrained by $E(f(x)/A) \geq 1/2$; hence there is a set $A^* \in \mathcal{A}$ that maximizes $\mu(A)$.

Note that in generic case we will have:

$$(m/k)\mu(B) < \mu(A^*) < ((m+1)/k)\mu(B) \quad \text{for some } m \in \{1, ..., k\} \; , \tag{6}$$

which means that since we will further need to "adjust" the set A^* to meet the districts' equality-of-size requirement (see below), the procedure in consideration may actually yield a set $\widetilde{A} \subseteq A^*$ such that $E(f(x)/\widetilde{A}) > 1/2$.

The only troublesome situation we may encounter is when

$$\mu(A^*) = (m/k)\mu(B) \quad \text{for some } m \in \{1, ..., k\} \tag{7}$$

and A^* maximizes $\mu(A)$ over the set $\mathcal{A} = \{A \subseteq B \mid E(f(x)/A) \geq 1/2\}$. In this case, such a districting procedure will, in expectation, produce a tie, not a victory over m districts; however, this knife-edge situation is not typical and we ignore it here.

1.2) Find m - the number of winning districts.

To satisfy the district's equal size requirement, we need to solve the following equation for $\tilde{k}_1$:

$$\mu(A^*)/\tilde{m} = \mu(B \setminus A^*)/(k - \tilde{m}) \tag{8}$$

Since k_1 is an integer, the number of winning districts equals the integer part of $\tilde{m}$: $m = i(\tilde{m})$. Note that if $\mu(A) < (1/k)\mu(B)$, then the number of winning districts will be zero. The higher the expectation of $f(x)$ over B and the larger k is, the easier it is to secure more winning districts.

1.3) Find the winning territory $\widetilde{A}$.

Define the set $\tilde{\mathcal{A}} = \{\widetilde{A} \subseteq A^* \mid \mu(\widetilde{A})/m = \mu(B \setminus \widetilde{A})/(k-m)\}$, which is non-empty by the corollary to the Lyapunov theorem (Dubins and Spanier). Choose $\widetilde{A} \in \tilde{\mathcal{A}}$ that maximizes $E(f(x)/\widetilde{A})$.

This procedure corresponds to "cutting off" some parts of A^* to meet the districts' equality of size requirement. By assumption 3 the party would want to cut off the regions of A^* with a low expected share of votes to increase the expected vote share on the rest of A^*, which will increase its probability of winning over this territory. Let $E(f(x)/\widetilde{A}) = \beta$, and observe that $\beta > E(f(x)/A^*) \geq 1/2$.

Thus we have defined $\widetilde{A}$, the territory that is to be further divided into k_1 parts to form the winning districts. To simplify the notation, we rename $\widetilde{A}$ as A. This concludes step 1 of the proof.

STEP 2. Given $A \subset R^2$, $f(x)$ such that $0 \le f(x) \le 1$ for any $x \in A$, $E(f(x)/A) > 1/2$, prove that there exists a partition $A_1, .., A_m$ of A such that the following properties are satisfied:

$$(i) \quad \mu(A_1) = ... = \mu(A_m) = (1/m)\mu(A) \ , \tag{9}$$
$$(ii) \quad E(f(x)/A_1) = ... = E(f(x)/A_m) = E(f(x)/A) \ , \tag{10}$$

where, as before, μ is the measure of a set's population.

To prove the above statement, we reproduce the reasoning used in the proof of lemma 1. For any set $C \subseteq A$ define a measure $\eta \in R^2$ by $\eta_1(C) = \mu(C)$, and $\eta_2(C) = \int_C f(x)d\mu$. Then by the Lyapunov's Theorem the set

$$R = \{(\eta_1(C), \eta_2(C)) \in R^2 \mid C \in \alpha\} \ ,$$

where α is the σ-algebra of A, is closed and convex. Since $(0,0) = (\eta_1(\emptyset), \eta_2(\emptyset)) \in R$ and $(\mu(A), \int_A f d\mu) = (\eta_1(A), \eta_2(A)) \in R$, we can find a set $A_1 \subset A$ such that

$$\mu(A_1) = \eta_1(A_1) = (1/m)\eta_1(A) = (1/m)\mu(A)$$

and

$$\eta_2(A_1) = \int_{A_1} f(x)d\mu = (1/m) * \int_A f(x)d\mu \ .$$

Then, by an induction argument, there exists a partition $A_1, .., A_m$ such that

$$\mu(A_i) = \eta_1(A_i) = (1/m)\eta_1(A) = (1/m)\mu(A) \ , \quad i = 1, .., m \ , \tag{11}$$

and

$$\eta_2(A_i) = \int_{A_i} f(x)d\mu = (1/m) * \int_A f(x)d\mu \ , \quad i = 1, .., m \ . \tag{12}$$

Now consider the expected vote share in each A_i:

$$\begin{aligned} E(f(x)/A_i) &= [\int_{A_i} f(x)d\mu]/\mu(A_i) = [(1/m) * \int_A f(x)d\mu]/[(1/m) * \mu(A)] = \\ &= [\int_A f(x)d\mu]/[\mu(A)] = E(f(x)/A) \ , \quad i = 1, .., m \ . \end{aligned} \tag{13}$$

Thus we proved that the partition of A we are looking for exists, which implies, using lemma 1, that there exists an optimal way to divide the "winning" territory into m winning districts. This completes step 2 of the proof.

Finally, to define the "losing" districts, we need to divide the rest of the territory, $B \setminus A$, into $(k - m)$ parts meeting the population equality requirement. By the corollary to the Lyapunov theorem (Dubins and Spanier), this is implementable.

Q.E.D.

Before turning to the question of the districts' connectedness, we consider whether imposing additional constraints on districting makes winning harder for a strategic party. The procedure that we have proposed allows us to predict the effect of additional equality constraints. Specifically, suppose now that the districts are required to be identical with respect to population size and also n other measurable characteristics such as ethnic composition. Then generally the proposed procedure would not work: by the corollary to Lyapunov theorem (Dubins and Spanier), it is possible to partition any set $A \subseteq B$ such that$f(x/A) \geq 1/2$ into m subsets $A_1, .., A_m$ in a way that all $(n+1)$ equality constraints will hold, but we cannot in a general secure that the additional equality constraints will hold between the subsets in A, $\{A_1, .., A_m\} \subseteq A$ and the subsets in $(B \setminus A)$, $\{A_{m+1}, .., A_k\} \subseteq (B \setminus A)$. However, in some special cases where distributions of different characteristics of a population are highly correlated with each other, the proposed procedure might be implementable. (Consider a degenerate case when all the characteristics are distributed identically and are perfectly correlated; then additional equality constraints do not matter at all!). The higher the number of equality constraints imposed, the less likely it is that the proposed procedure is implementable. Thus additional equality constraints imposed on the districts make it generally harder to manipulate electoral outcomes.

What can be said about other types of constraints that may be imposed on districting? Although a formal analysis of the effect of all possible constraints is not the subject of this paper, we can note that generally additional constraints restrict the freedom of a decision-maker and thus make gerrymandering harder. On the other hand, the same argument shows that too many constraints might prevent "fair" districting as well as gerrymandering, and hence we cannot blindly welcome more restrictions instead of fewer. Yet in some situations constraints such as, for example, the requirements not to split ethnic communities or counties (if we treat them as "fairness" constraints) might restrict gerrymandering. As an illustration, suppose that all voters in the same ethnic community are either supporters or opponents of the gerrymandering party. Then the "not splitting" requirement will prevent the party from spreading its supporters among numerous districts, as it would otherwise choose to do. Unfortunately, we cannot generalize this conclusion for an arbitrary distribution of voters over counties or ethnic groups. Rather we can conclude that effectiveness of each additional constraint will depend on characteristics of voters' distribution over the territory.

We now go back to our initial problem. Proposition 1 assures that there is a solution to the districting problem when the districts are not required to be connected. In lemma 2 below, we show that the disconnected parts of the districts may be further connected in a way that nearly preserves the required characteristics of districts. The idea is to connect all the disconnected parts with connecting sets of infinitesimal measure of population size[4]. Then by continuity arguement we can further show that the expected share of votes does not change significantly in any district, either.

Lemma 2 *For every optimal map*[5] $A_1, .., A_k$, *for any* $\delta > 0$, *there exists a connected map* $F_1, .., F_k$ *such that*

1. *Each* F_i *is connected,* $i = 1, .., k$;
2. $\mid \mu(F_i) - \mu(A_i) \mid < \delta$ *for all* $i = 1, .., k$.

Outline of the Proof of Lemma 2 Suppose that the procedure described in Proposition 1 has been implemented and we obtained an optimal partition of the set B into the districts $A_1, .., A_k$, with each A_i possibly consisting of a number of disjoint parts (subsets) A_{in}. To make connections possible, it is necessary for the following conditions to be satisfied: (i) Every A_i has a finite boundary. Since Proposition 1 does not provide any information about the properties of the resulting districts of the optimal map, we cannot exclude the case that some of the sets (districts) might have infinite boundaries[6]; this may create difficulties in connecting the disjoint parts. (ii) No part of any district forms a loop, i.e. every A_{in} is contractible. Otherwise it might be impossible to connect disjoint parts of a district together without cutting some other district into pieces. (iii) No resulting connected district $\tilde{A}_i$ forms a loop, i.e. $\tilde{A}_i$ is contractible. These conditions, as we show in the proof, guarantee that it is possible to connect each of k districts together with no district cutting any other into parts; the connecting sets can be located along the boundaries. Hence we first adjust each set A_i to meet the above conditions; then we construct the connecting sets. By making the adjustments and the connecting sets small enough we guarantee that the total measure of population does not change significantly in the districts after the adjustments are made and the connecting parts are added.

Proof of Lemma 2 By Proposition 1, there exist an optimal districting map $A_1, .., A_k$ when the connectedness requirement is not imposed. We take this map as a starting

[4]Note that this in fact has been done in practice, at least in some racial gerrymandering cases. I am indebted to Morgan Kousser for pointing out the following examples. *Gomillion v. Lightfoot* (364 U.S. 339) Supreme Court case (1960) considered validity of redefining city boundaries in a way that removed the black voters from its territory; the new map of the city contained parts that were connected solely by a channel. In California, Proposition 6, passed in June, 1980, was aimed at preventing just such tricks (among other things). Yet a congressional district set up in 1981 was connected only by water (see district 6 in the Congressional district maps for the 1980s.).

[5]The result of this lemma holds for an arbitrary, and not only an optimal or a feasible, map.

[6]Although a district with an infinite boundary is hard to imagine, Lyapunov theorem does not guarantee that the resulting subsets (districts) will always have "nice" properties, such as finite boundaries. Hence we need to undertake some additional steps to make sure that we can implement the districting procedure in question.

point for costruction of connected districts. Now suppose given is a partition of the set B into subsets $A_1, .., A_k$, $B = \bigcup_{i=1}^{k} A_i$, with each of A_i possibly consisting of a number of disjoint parts (subsets) $A_i = \bigcup_{n=1}^{N_i} A_{in}$, and $A_{in}^{\circ} \cap A_{jm}^{\circ} = \emptyset$ for $i \neq j$ or $n \neq m$, $n = 1, .., N_i$, $m = 1, .., N_j$, $N_i, N_j \leq \infty$, $i, j = 1, .., k$.

STEP 1 Adjust each set A_i to make the boundaries finite.

1.1) Approximate each A_i with a finite collection of rectangles. Observe that since R^2 is a metric space we can cover every A_i with an infinite collection of rectangles: $A_i \subset \bigcup_{m=1}^{\infty} R_{im}$ for all $i = 1, .., k$. Moreover, for any $\epsilon > 0$ we can find $\{R_{im}\}_m$ and numbers M_i such that

$$\mu\{[(\bigcup_{m=1}^{M_i} R_{im}) \setminus A_i] \cup [A_i \setminus (\bigcup_{m=1}^{M_i} R_{im})]\} < \epsilon/2k \text{ for all } i, \tag{14}$$

which implies

$$|\mu(\bigcup_{m=1}^{M_i} R_{im}) - \mu(A_i)| < \epsilon/2k \text{ for all } i. \tag{15}$$

Thus we approximate every A_i with a finite collection of rectangles in a way that does not change the measure of the set significantly; this construction guarantees that every resulting set $S_i = \bigcup_{m=1}^{M_i} R_{im}$ has a boundary of a finite length.

FIGURE 1 HERE

1.2) Estimate the size of overlapping and remaining parts. As a result of the above adjustment some of S_i's may overlap, and some parts of A_i's may be left uncovered by any of S_i's (we will further call the latter parts the "remaining" sets). We now show that the overlapping and remaining parts are negligibly small in size and hence can be added to any district without changing its characteristics significantly.

(i) Prove that the the overlapping parts are negligibly small. It is sufficient to show that

$$\mu(S_i \setminus \bigcup_{j \neq i} S_j) \rightarrow \mu(S_i) \text{ as } \epsilon \rightarrow 0 \text{ for all } i = 1, .., k \; . \tag{16}$$

Let $e_i = (S_i \setminus A_i)$ and $r_i = (A_i \setminus S_i)$. Then by construction $S_i = (A_i \setminus r_i) \cup e_i$ and from (14)

$$\mu(e_i) + \mu(r_i) = \mu[(S_i \setminus A_i) \cup (A_i \setminus S_i)] < \epsilon/2k \; ,$$

which implies that for all i

$$\mu(e_i) < \epsilon/2k \text{ and } \mu(r_i) < \epsilon/2k \ . \tag{17}$$

If S_i and S_j overlap then for any $i, j = 1, .., k$, $i \neq j$,

$$\begin{aligned} \mu[S_i \cap S_j] &= \mu\{[(A_i \setminus r_i) \cup e_i] \cap [(A_j \setminus r_j) \cup e_j]\} = \\ &= \mu\{[(A_i \setminus r_i) \cap (A_j \setminus r_j)] \cup [(A_i \setminus r_i) \cap e_j] \cup [(A_j \setminus r_j) \cap e_i] \cup [e_i \cap e_j]\} = \\ &= \mu[(A_i \setminus r_i) \cap (A_j \setminus r_j)] + \mu[(A_i \setminus r_i) \cap e_j] + \mu[(A_j \setminus r_j) \cap e_i] + \mu[e_i \cap e_j]\} = \\ &= \mu[(A_i \setminus r_i) \cap e_j] + \mu[(A_j \setminus r_j) \cap e_i] + \mu[e_i \cap e_j]\} \end{aligned}$$

since $A_i^\circ \cap A_j^\circ = \emptyset$. It follows that

$$\mu(S_i \cap S_j) \leq \mu(e_i \cup e_j) = \mu(e_i) + \mu(e_j) < 2\epsilon/2k \ . \tag{18}$$

FIGURE 2 HERE

Similarly we can estimate the total overlapping of all S_i's:

$$\mu[\bigcup_{i,j,\, i \neq j} (S_i \cap S_j)] \leq \mu(\bigcup_{i=1}^{k} e_i) \leq \sum_{i=1}^{k} \mu(e_i) < k * (\epsilon/2k) = \epsilon/2 \ . \tag{19}$$

Hence the total size of the overlapping districts can be made arbitrary small. Moreover, since

$$(S_i \setminus \bigcup_{j \neq i} S_j) = S_i \setminus \bigcup_{j,k,\, j \neq k} (S_j \cap S_k) \ ,$$

we obtain

$$\begin{aligned} \mu(S_i \setminus \bigcup_{j \neq i} S_j) &= \mu(S_i \setminus \bigcup_{j,k,\, j \neq k} (S_j \cap S_k)) = \mu(S_i) - \mu(\bigcup_{j,k,\, j \neq k} (S_j \cap S_k)) > \\ &> \mu(S_i) - \epsilon/2 \to \mu(S_i) \text{ as } \epsilon \to 0 \ , \end{aligned} \tag{20}$$

which proves (16).

(ii) Estimate the size of the remaining sets. Consider the colection of remaining sets $\{r_I\}_I$, where r_I denotes the remaing set adjacent to the sets S_{im}, $\{im\} \in I$, $I \subseteq$

$\{(ij)\}_{i=1,..,k}^{j=1,..,N_i}$. For example, a set $r_{\{im,jn\}}$ would be the set left uncovered between S_{im} and S_{jn}. By construction

$$\begin{aligned} r_I \ &\subseteq \ (\bigcup_{\{im\}\in I} A_{im}) \setminus (\bigcup_{\{im\}\in I} S_{im}) = \bigcup_{\{im\}\in I} (A_{im} \setminus (\bigcup_{\{im\}\in I} S_{im})) \subseteq \\ &\subseteq \ \bigcup_{\{im\}\in I} (A_{im} \setminus S_{im}) \subseteq \bigcup_{i:\{im\}\in I} (A_i \setminus S_i) \subseteq \bigcup_{i:\{im\}\in I} r_i \ , \end{aligned} \tag{21}$$

and hence

$$\bigcup_I r_I \subseteq \bigcup_i r_i \ . \tag{22}$$

Now consider the size of all overlapping and"remaining" sets. Since from (19) and (22)

$$[\bigcup_I r_I] \cup [\bigcup_{j,k,\, j\neq k} (S_j \cap S_k)] \subseteq [\bigcup_i r_i] \cup [\bigcup_i e_i] = \bigcup_i (r_i \cup e_i) \ , \tag{23}$$

we have

$$\begin{aligned} \mu\{[\bigcup_I r_I] \cup [\bigcup_{j,k,\, j\neq k} (S_j \cap S_k)]\} \leq \mu[\bigcup_i (r_i \cup e_i)] \leq \\ \leq \sum_{i=1}^{k} \mu(r_i \cup e_i) < k * (\epsilon/2k) = \epsilon/2 \ , \end{aligned} \tag{24}$$

which proves that the overlapping and the remaining sets are very small and cannot affect characteristics of any set they will be added to substantially.

1.3) Construct the sets B_i with finite boundaries and such that $\bigcup_i \overline{B}_i = B$.

First note that by construction each $(S_i \cap S_j)$ is a finite collection of rectangles; so is each r_I which is surrounded by S_i's from all sides; r_I's with the parts adjacent to the boundary of B might not wholy consist of rectangles in this part but since we can assume that the boundary of B is finite, r_I's boundaries are guaranteed to be finite as well.

It follows that we can construct the sets B_i with finite boundaries in the following way: start with the sets $(S_i \setminus \bigcup_{j\neq i} S_j)$; then arbitrarily add each $(S_i \cap S_j)$ to either $(S_i \setminus \bigcup_{j\neq i} S_j)$ or $(S_j \setminus \bigcup_{i\neq j} S_i)$ for all $i = 1,..,k$; then arbitrarily add each r_I to one of the adjacent S_{im}'s: $(im) \in I$. Then by construction

$$\begin{aligned} (S_i \setminus \bigcup_{j\neq i} S_j) \ &\subseteq \ B_i \ \subseteq \ S_i \cup (\bigcup_{j,k,\, j\neq k} (S_j \cap S_k) \cup (\bigcup_i r_i) \ , \\ \mu(S_i) - \epsilon/2 \ &\leq \ \mu(B_i) \ \leq \ \mu(S_i) + \epsilon/2 \ , \\ |\mu(B_i) - \mu(S_i)| \ &\leq \ \epsilon/2 \ ; \end{aligned} \tag{25}$$

and since from (19) we also have

$$| \mu(S_i) - \mu(A_i) | \leq \epsilon/2k \ ,$$

we finally obtain

$$| \mu(B_i) - \mu(A_i) | < \epsilon \ . \tag{26}$$

Thus we have constructed the sets B_i , $i = 1,..k$, which cover the whole territory B, have finite boundaries and differ in size from the original sets A_i by arbitrarily small numbers. Next, for the purposes of further construction, make every B_i open by excluding its boundaries. Formally, for every $i = 1,..,k$, take $\widehat{B_i} = (\overline{B}_i)^\circ$. We now have every district represented by an open set $\widehat{B_i} = \bigcup_{n=1}^{N_i} \widehat{B}_{in}$, with population measure in all the districts unchanged: $\mu(\widehat{B_i}) = \mu(B_i)$. Besides note that now $N_i < \infty$ for all i since every $\widehat{B_i}$ is a union of a finite number of rectangles.

STEP 2 Make all the sets $\widehat{B}_{in}$, $i = 1,..,k$, $n = 1,..,N_i$, contractible. In case some of $\widehat{B}_{in}$'s form loops around other sets, we need to break these loops with the connecting sets C_j of measure zero, $\mu(C_j) = 0$[7], to make it possible to connect subsets inside and outside loops. Since each C_j can be drawn along the borders of rectangles R_{im}, we can guarantee that all C_j's will be of finite lengths. Thus C_j's are the additional boundaries created to make every subset $\widehat{B}_{in}$ contractible, and the measure of any $\widehat{B}_i$ has not changed.

Denote by $\{C_j\}_{j=1}^{L}$ the collection of all the boundaries in B, where L is some number, $L < \infty$, and each C_j denotes a piece of boundary of finite length which either cuts the loop, $C_j = \widehat{B}_{in} \setminus (\overline{\widehat{B}}_{in})^\circ$, or separates two neighboring subsets, $C_j = \overline{\widehat{B}}_{in} \cap \overline{\widehat{B}}_{hm}$, for some $i \neq h$ or $n \neq m$. Then $\bigcup_{j=1}^{L} C_j = B \setminus (\bigcup_{i=1}^{k} \bigcup_{n=1}^{N_i} \widehat{B}_{in})$, where $\widehat{B}_{in}$'s are now contractible sets with finite boundaries, and $\mu(B) = \mu(\bigcup_{i=1}^{k} \bigcup_{n=1}^{N_i} \widehat{B}_{in}) = 1$, $\mu(C_j) = 0$ for every j.

STEP 3 For every district i, connect its disjoint parts. Fix an $i \in \{1,..k\}$. Take a minimal – with respect to the number of boundaries C_j included – set $D_i = (\bigcup_{n=1}^{N_i} \widehat{B}_{in}) \cup (\bigcup_j C_j)$ such that it is connected. Minimality guarantees the absence of loops in the resulting set. Denote by $\{C_{ij}\}_{j=1}^{L_i}$ the collection of boundaries included in D_i. Next, to give the resulting connected district a positive measure of population size in every point, we need to cover all the C_{ij}'s with sets of arbitrary small but non-zero measures. Since every C_{ij} is of finite length, for any $\epsilon > 0$ and for every C_{ij}, $j = 1,..,L_i$, we can find a finite open cover $E_{ij} = \bigcup_h E_{ijh}$ such that each E_{ijh} is a rectangle, $C_{ij} \subset \bigcup_h E_{ijh}$ and $\mu(\bigcup_h E_{ijh}) < \epsilon/(k * L_i)$. The open set $F_i = (\overline{(\bigcup_{n=1}^{N_i} \widehat{B}_{in}) \cup (\bigcup_j \bigcup_h E_{ijh})})^\circ$ is now fully connected, and its population measure does not exceed the measure of the initial set A_i by more than $2\epsilon * \mu(A_i)$.

FIGURE 3 HERE

[7] This simply means that C_j's are lines in R^2.

The same argument applies to show that the population measure on the sets from which the connecting sets E_{ij} are cut does not change significantly.

We have now obtained an open connected set F_i which is equivalent to the initial disconnected set A_i up to an arbitrarily small deviation in population measure. The construction of F_i caused the change of the boundaries but since F_i is an open union of rectangles the new finite boundaries have automatically emerged, and the set of all boundaries $\{C_j\}_{j=1}^{L}$ in B is once again well defined.

Repeat the procedure described in Step 3 for all $i \in \{1,..k\}$. Observe that the total number of connecting sets E_{ij}, $i = 1,..,k$, $j = 1,..,L_i$, is finite and hence the deviations from original values of population size measure can be kept arbitrary small for every district.

Q.E.D.

It is left to show that, once the districts are connected, every winning districts is still winning, i.e. the expectation of votes in every winning district remains on the level above one half. We use the following lemma to show that negligible changes in the population size of a set can result in only small changes in the expected share of votes over this set[8].

Lemma 3 *If $f(x)$ is a continuous function on a measurable space $(B, \mathcal{B})$, then the expectation of votes $E[f(x)/C]$ on any $C \in \mathcal{B}$ is continuous in the measure of population size $\mu(C)$: for any $C, D \in \mathcal{B}$, any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that if $\mu[(C \setminus D) \cup (D \setminus C)] < \delta$ then $E[f(x)/C] - E[f(x)/D] < \epsilon$. In particular, any $\delta \leq [\epsilon * \min\{\mu(C), \mu(D)\}]/[2 * \max_{x \in B} f(x)]$ satisfies the the above requirment.*

We are now in position to present the main result of the paper.

Proposition 2 *For any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that the territory B can be divided into k connected districts of δ-close to equal size and ϵ-close in the expected vote share to the optimal map. In particular, generically, the number of winning districts in the resulting connected map stays the same as in the optimal map.*

Proof of Proposition 2 The first statement of the proposition follows directly from lemmas 2 and 3. Therefore, it is sufficient to describe the procedure that produces a connected map with the same number of winning districts as in an optimal map. Take an arbitrarily optimal map $M^* = \{A_1, .., A_k\}$. Then by the properties of the optimal

[8]The proof of lemma 3 is given in the appendix.

map (lemma 1) $E[f(x)/A_i] = E[f(x)/A_j]$ for any two winning districts A_i, A_j. Let ϵ_1 denote the value

$$\epsilon(M^*) = E[f(x)/A_i] - 1/2 \quad \text{for some} \quad A_i \in M^* \ . \tag{27}$$

Then, by lemma 3, and keeping in mind that $\mu(A_i) = 1/k$ for any i, define $\delta(\epsilon(M^*))$ by

$$\delta(\epsilon(M^*)) = \frac{\epsilon(M^*)}{k * (2 * \max_{x \in B} f(x) + \epsilon(M^*))} \tag{28}$$

By lemma 2, for any given $\delta > 0$ there exists a connected map $F = \{F_1, .., F_k\}$ such that $|\mu(A_i) - \mu(F_i)| < \delta/2$ for all $i = 1, .., k$. Then by lemma 3 we have that if $|\mu(A_i) - \mu(F_i)| < \delta(\epsilon(M^*))/2$, then

$$|\mu(F_i) - \mu(F_j)| < \delta(\epsilon(M^*)) \quad \text{for every} \quad i, j = 1, .., k \tag{29}$$

and

$$E[f(x)/F_i] > E[f(x)/A_i] - \epsilon(M^*) = 1/2 \quad \text{for all} \quad i \in I(M^*) \ . \tag{30}$$

Thus for any $\delta < \delta(\epsilon(M^*))/2$ by lemma 2 there exists a connected map such that

$$|\mu(F_i) - \mu(F_j)| < \delta \tag{31}$$

and, for any $i = 1, .., k$,

$$\text{if} \quad E[f(x)/A_i] > 1/2 \quad \text{then} \quad E[f(x)/F_i] > 1/2 \ . \tag{32}$$

Q.E.D.

We conclude that generically the party can insure for itself the same number of winning districts even if the additional connectedness requirement is imposed.

Next we present a simple example of the suggested districting procedure.

3 The case of a single-peaked distribution of the expected share of votes

Suppose the voters are uniformly distributed over the territory $B \subset R^2$ and the distribution of the expected share of votes $f(x)$ that the party will get in the elections is single-peaked with its peak at some point O in the interior of B; and suppose $f(x)$ is decreasing in all directions from O. Assume that $f(x)$ is continuous at any point x of the territory, and $0 \leq f(x) \leq 1$. How should the party divide the territory into k districts of equal size in this case?

Given the assumption about the distribution of $f(x)$, we might expect that the winning districts will be concentrated around the center, so that each of them will cover some central area with a high expectation of the vote share and then expand towards the border as long the expectation of the vote share in the district exceeds one half. More accurately, the party will want to find the largest possible neighborhood A of O in B such that the expected share of the votes all over this neighborhood slightly exceeds one half, $E(f(x)/A) > 1/2$. This region will then be divided into k_1 winning districts, and the rest of the territory – distant from the peak O – will constitute k_2 losing districts, $k_1 + k_2 = k$.

For simplicity, suppose that the territory B is a circle with the center at O and the radius R (denote it by $B(O; R)$), and the expected share of the votes is maximal at O, with $f(O) > 1/2$. Let us consider the requirements on the distribution $f(x)$ of expected vote shares that will allow the party to get the majority of seats in elections under these circumstances.

Suppose k is even. Then at least $(k/2 + 1)$ winning districts are necessary to win in the election. Following the reasoning presented above, without loss of generality we can assume that the winning districts may be drawn as k_1 identical sectors of a circle $A(O; r)$ with the center at O and with radius r, where $r \leq R$ is determined to satisfy the condition that the expected share of the votes over the area of each sector of A is no less than one half:

$$E(f(x)/A_i) \geq 1/2 \ , \ i = 1, .., k_i \ ,$$

where A_i denotes the i-th sector of A. Equivalently,

$$[\int_0^r (1/k_1)\pi x^2 f(x) dx]/(1/k_1)\pi x^2 \geq 1/2 \ ,$$

or:

$$[\int_0^r x^2 f(x) dx]/x^2 \geq 1/2 \ . \tag{33}$$

To determine the number of winning districts, the party solves the following equation for k_1:

$$\pi r^2/k_1 = \pi(R^2 - r^2)/(k - k_1) \tag{34}$$

Equation (34) is the equal size of districts requirement. If the k_1 that solves this equation is not an integer, then its integer part will indicate the number of winning districts. Note that in the latter case r will need to be adjusted – namely, decreased to satisfy (33).

We can now write out the party's problem as follows:

$$\begin{aligned}
&\max k_1 \\
&\text{subject to:} \\
&k_1 \in \{1, 2, .., k\} \\
&\left(\int_0^r \pi x^2 f(x)dx\right)/\pi x^2 > 1/2 \\
&\pi r^2/k_1 = \pi(R^2 - r^2)/(k - k_1)
\end{aligned}$$

The party may expect to win the majority of seats in a legislative body if $k_1 \geq k/2+1$, that is the number of winning districts will be greater than a half of the total number of districts. In particular, it is sufficient to have $k = k/2 + 1$, and thus it is sufficient to find r which solves:

$$\left(\int_0^r \pi x^2 f(x)dx\right)/\pi x^2 > 1/2 \tag{35}$$

and

$$\pi r^2/(k/2+1) = \pi(R^2 - r^2)/(k/2-1) \ . \tag{36}$$

Obviously, the higher the expected value of $f(x)$, the more easily conditions (35) and (36) are satisfied. It is interesting to note that under these conditions, to win a majority of districts it is sufficient to have slightly more than half a territory with slightly more than half of the supporters. The total share of supporters over the whole territory, or, equivalently, the expected share of voters for the party necessary to win the election may be much lower than one half. What is important for winning, though, is having a non-uniform distribution of supporters over the territory, with a concentration of supporters in some areas. Then the party will draw the districting map in a way that will secure it a maximal number of winning districts located around the peak of the distribution of expected vote shares. In Figure 4, we illustrate the districting map problem for the case when the number of districts k equals seven.

FIGURE 4 HERE

4 Conclusions

In this paper we show that for any given territory it is theoretically possible to find a districting map that will in expectation maximize the number of winning districts for a gerrymandering party, under the condition that no restrictions except for the equal population of districts requirement are imposed. If the districts are required to have connected territories, optimal districting is still implementable if the districts are allowed to vary in population size by an arbitrarily small amount. With the optimal map the political party in control of the districting process might need much less than one half of the electorate's support to be able to win in a majority of seats in a legislative body. The party will choose to concentrate its supporters in the group of winning districts while leaving the other group of losing districts to the party's competitors. These findings are consistent with earlier results of Owen and Grofman (1988); we show that they hold not only for an "ideal", but also for a geographically constrained optimal districting scheme.

Although we have only considered the case of a party which takes expecteed vote share as a perfect predictor of electoral outcomes, and in in elections held under simple majority rule, the approach presented can be generalized for any m-majority rule[9] or for a party with any attitude towards risk. For example, if the party is risk-averse or is aware of possible shifts in electoral votes, it might want to secure the expected share of the votes in the winning districts at some level higher than one half depending on the degree of the risk-aversion. The trade-off that the party faces then is between the number of potential winning districts and the risk of not winning all of them. The procedure proposed here does not substantially change in this case except for the value of the expected share of votes that the party would want to secure in the winning districts.

We are now in position to return to the question we asked in the introduction: may the constraints on districting help to prevent gerrymandering? As we find in the paper, imposition of the connectedness requirement generically does not provide an effective controlling device against gerymandering: the party is able to secure for itself the same number of winning districts as in the unrestricted optimal scheme. However, additional requirements imposed on districts make manipulability of electoral outcomes by means of districting more difficult. Specifically, we find that the optimal gerrymandering procedure considered above may not be extended to an arbitrary number of districting constraints. Thus by imposing extra requirements, either on the population characteristics or shapes of districts, an electorate might at least get more protection against possible partisan gerrymandering (if not more fairness).

Consideration of geographic constraints, when the distribution of population and partisan support is exogenously given, helps to explain why districts of unusual shapes, extended too much in one direction or with narrow connecting parts included, might present evidence of a partisan gerrymander. Thus we find some rational grounds for most people's understanding of gerrymandering as drawing oddly shaped districts (Butler and Cain, 1992). Yet one should not forget that some districts may look "funny" just because they join what some redistricters consider communities of interest, or because the topology of an area is not uniform.

Although it is hard to believe that the party in control of the districting process would use an abstract optimization scheme like the one we have presented, one might expect to find certain real world approximations of the proposed procedure. Empirical

[9] By an m-majority rule we mean a rule requiring an m-percent majority of votes to win a district.

consideration of electoral maps in view of our theoretical scheme suggests a subject for an exciting but still a different study.

Appendix

Proof of Lemma 1 It is sufficient to prove that if $A_i \in M^*$ – optimal and i solves

$$\min_{i \in I(M^*)} E(f(x)/A_i) \ , \tag{37}$$

then $E(f(x)/A_j) = E(f(x)/A_i)$ for all $j \in I(M^*)$. Suppose there exists A_j, $j \in I(M^*)$, such that $E(f(x)/A_i) < E(f(x)/A_j)$. Then for any set $C \subseteq (A_i \cup A_j)$ define a measure $\eta \in R^2$ by $\eta_1(C) = \mu(C)$, and $\eta_2(C) = \int_C f(x)d\mu$. Then by the Lyapunov's Theorem the set

$$R = \{(\eta_1(C), \eta_2(C)) \in R^2 \mid C \in \alpha\} \ ,$$

where α is the σ-algebra of $A_i \cup A_j$, is closed and convex. Since $(0,0) = (\eta_1(\emptyset), \eta_2(\emptyset)) \in R$ and $(\mu(A_i \cup A_j), \int_{A_i \cup A_j} f d\mu) = (\eta_1(A_i \cup A_j), \eta_2(A_i \cup A_j)) \in R$, we can find a a partitition of $A_i \cup A_j$ into two subsets $\tilde{A}_i$, $\tilde{A}_j$ such that

$$\eta_1(\tilde{A}_i) = \eta_1(\tilde{A}_j) = (1/2)\eta_1(A_i \cup A_j) = \mu(A_i) = \mu(A_j)$$

and

$$\eta_2(\tilde{A}_i) = \int_{\tilde{A}_i} f(x)d\mu = (1/2) * \int_{A_i \cup A_j} f(x)d\mu \ ,$$

$$\eta_2(\tilde{A}_j) = \int_{\tilde{A}_j} f(x)d\mu = (1/2) * \int_{A_i \cup A_j} f(x)d\mu \ .$$

Now compare the expected vote share in A_i with the ones in $\tilde{A}_i$ and $\tilde{A}_j$:

$$E(f(x)/(A_i \cup A_j)) = [E(f(x)/A_i) + E(f(x)/A_j))]/2 > E(f(x)/A_i) \tag{38}$$

by the initial assumption;

$$\begin{aligned} E(f(x)/(\tilde{A}_i)) &= [\int_{\tilde{A}_i} f(x)d\mu]/\mu(\tilde{A}_i) = \\ &= [(1/2) * \int_{A_i \cup A_j} f(x)d\mu]/[(1/2) * \mu(A_i \cup A_j)] = \\ &= [\int_{A_i \cup A_j} f(x)d\mu]/[\mu(A_i \cup A_j)] = \\ &= E(f(x)/(A_i \cup A_j)) \ , \end{aligned} \tag{39}$$

and, similarly,

$$E(f(x)/(\tilde{A}_j)) = E(f(x)/(A_i \cup A_j)) \ . \tag{40}$$

Thus, there exists another map $\tilde{A} = (M^* \setminus (A_i \cup A_j)) \cup (\tilde{A}_i \cup \tilde{A}_j))$ that is prefered to M^* by assumption 3. This shows that the initial map M^* could not have been optimal.

Q.E.D.

Proof of Lemma 3 Suppose we have $\mu[(C \setminus D) \cup (D \setminus C)] < \delta$ for some $\delta > 0$; this also implies $\mu(C \setminus D) < \delta$ and $\mu(D \setminus C) < \delta$; moreover,

$$\begin{aligned}
\mid \mu(C) - \mu(D) \mid &= \mid \mu[(C \setminus D) \cup (C \cap D)] - \mu[(D \setminus C) \cup (D \cap C)] \mid = \\
&= \mid \mu(C \setminus D) + \mu(C \cap D) - \mu(D \setminus C) - \mu(D \cap C)] \mid = \\
&= \mid \mu(C \setminus D) - \mu(D \setminus C)] \mid < \delta \ ,
\end{aligned}$$

and $\mu(C) - \mu(C \cap D) < \delta$, $\mu(D) - \mu(D \cap C) < \delta$. Now consider the expected share of votes on C and D:

$$\begin{aligned}
E[f(x)/C] &= [\int_C f(x) d\mu]/\mu(C) = \\
&= [\int_{(C \setminus D) \cup (C \cap D)} f(x) d\mu]/\mu[(C \setminus D) \cup (C \cap D)] \ = \\
&= [\int_{(C \setminus D)} f(x) d\mu]/\mu[(C \setminus D) + \mu(C \cap D)] \ + \\
&\quad + \ [\int_{(C \cap D)} f(x) d\mu]/\mu[(C \setminus D) + \mu(C \cap D)] \ = \\
&= E[f(x)/(C \setminus D)] * [\mu(C \setminus D)/\mu(C)] \ + \\
&\quad + \ E[f(x)/(C \cap D)] * [\mu(C \cap D)/\mu(C)] \ .
\end{aligned} \tag{41}$$

Similarly,

$$\begin{aligned}
E[f(x)/D] &= E[f(x)/(D \setminus C)] * [\mu(D \setminus C)/\mu(D)] \ + \\
&\quad + \ E[f(x)/(C \cap D)] * [\mu(C \cap D)/\mu(D)] \ .
\end{aligned} \tag{42}$$

Let $a = \max\{\mu(C), \mu(D)\}$, and $b = \min\{\mu(C), \mu(D)\}$. Then we obtain

$$\begin{aligned}
&E[f(x)/C] - E[f(x)/D] \leq \\
&\quad \leq \{\max_{x \in B} f(x)\} * \{\mu(C \setminus D)/\mu(C)\} - -\{\min_{x \in B} f(x)\} * \{\mu(D \setminus C)/\mu(D)\} + \\
&\quad +\{\max_{x \in B} f(x)\} * \{[\mu(C \cap D)/\mu(C)] - [\mu(D \cap C)/\mu(D)]\} < \\
&\quad < \{\max_{x \in B} f(x)\} * \{\delta/b\} + \{\max_{x \in B} f(x)\} * \{\delta/b\} = \\
&\quad = \{\max_{x \in B} f(x)\} * \{2\delta/b\} \ .
\end{aligned} \tag{43}$$

Hence for an arbitrary $\epsilon > 0$, if we choose

$$\delta(\epsilon) = [\epsilon * \min\{\mu(C), \mu(D)\}]/[2 * \max_{x \in B} f(x)] \tag{44}$$

then the continuity requirment stated in the lemma is satisfied.

Q.E.D.

References

[1] Balinski, Michel L., and H. Peyton Young. 1982. *Fair Representation.* Yale University Press, New Haven and London.

[2] Bridge, John R. 1983. "Redistricting to Maximize the Preservation of Political Boundaries." *Social Science Research* 12: 205-214.

[3] Butler, David, and Bruce Cain. 1992. *Congressional Redistricting: Comparative and Theoretical Perspectives.* Macmilian Publishing Company, New York.

[4] Cain, Bruce. 1984. *The Reappointment Puzzle.* University of California Press.

[5] Dubins, L. E. and E. H. Spanier. 1961. "How to Cut a Cake Fairly." *American Mathematic Monthly* 68: 437-459.

[6] Grofman, Bernard (ed.). 1990. *Political Gerrymandering and the Courts.* Agathon Press, New York.

[7] Grofman, Bernard, A. Lijphart, R. Mckay and H. Scarrow (ed.). 1982. *Representation and Redistricting Issues.* Lexington Books, Lexington.

[8] Hildenbrand, Werner. 1974. *Core and Equilibria of a Large Economy.* Princeton University Press, Princeton, New Jersey. Part 1.

[9] Owen. Guillermo, and Bernard Grofman. 1988. "Optimal Partisan Gerrymandering." *Political Geography Quarterly* 7, No. 1: 5-22.

[10] *Gomillion v. Lightfoot*, 364 U.S. 339 (1960). *81 Supreme Court Reporter.*

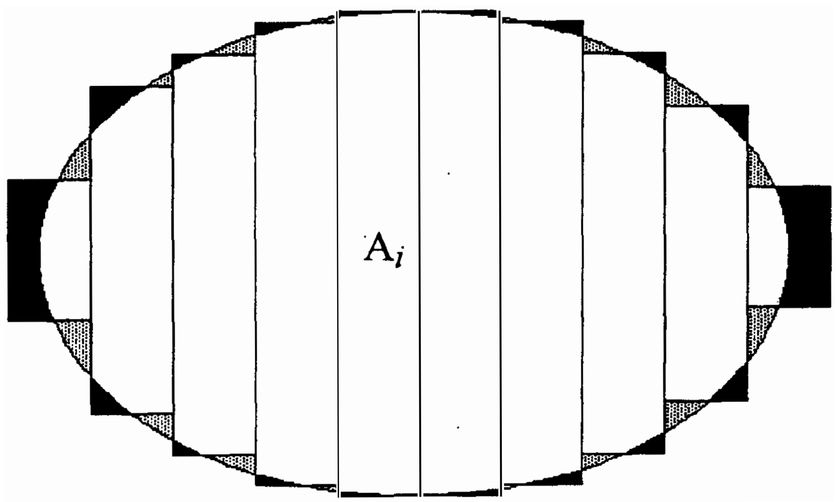

Figure 1. Approximating a set A_i with a finite collection of rectangles S_i

■ $e_i = S_i \setminus A_i$
■ $r_i = A_i \setminus S_i$

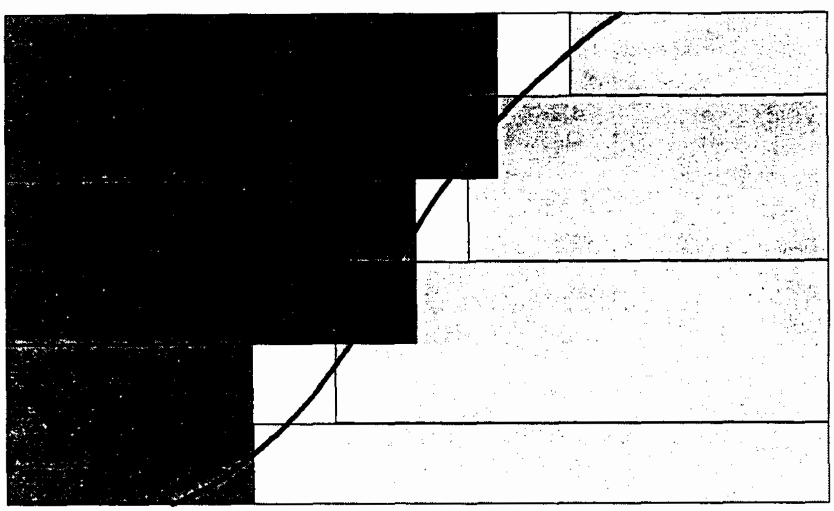

Figure 2. Overlapping and remaining sets:

boundary divining A_i and A_j
■ S_i
□ S_j
■ overlapping set $S_i \cap S_j$
□ remaining set $r_{\{i,j\}} = (A_i \cup A_j) \setminus (S_i \cup S_j)$

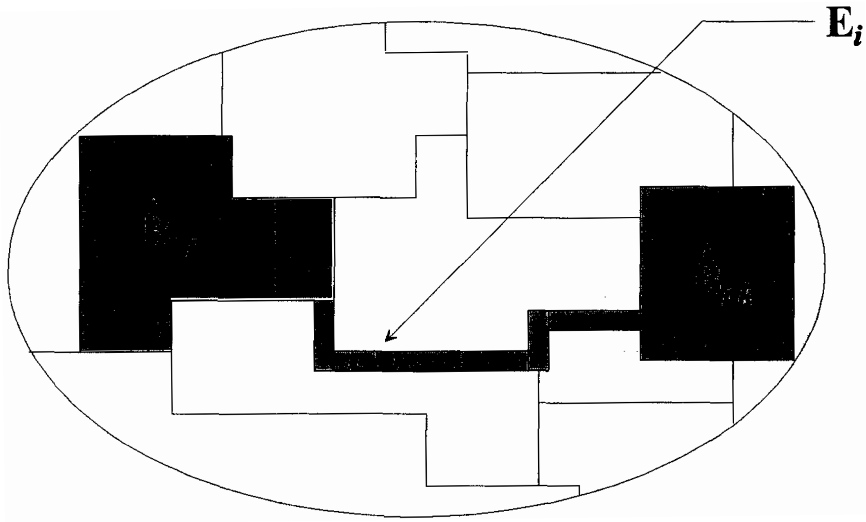

Figure 3. Connecting district's disjoint parts.

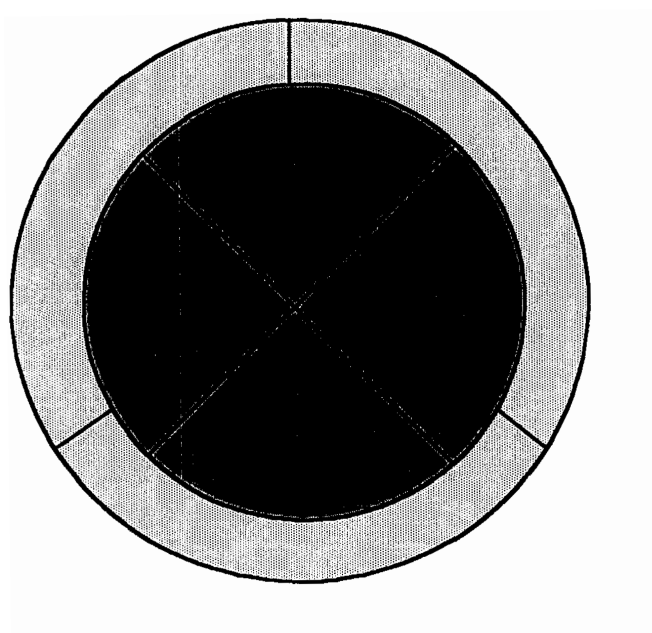

Figure 4. Dividing a circle into seven districts:

winning districts

losing districts